Claims

1. Method for automatically determining a characterizing strength (C) which indicates
   how well a text (11) stored in a database (10) describes a query (15), comprising the
   steps of:
   a) defining a query (15) comprising a query word;
   b) creating (71) a graph (30) with nodes and links, whereby words of the text (11) are
       represented by the nodes and a relationship between the words is represented by the
       links;
   c) evolving (72) the graph (30) according to a pre-defined set of rules,
   d) determining a neighborhood of the query word, the neighborhood comprising those
       nodes connected through one or more links to the query word; and,
   e) calculating the characterizing strength (C) based on the neighborhood.

2. The method of claim 1, wherein the characterizing strength (C) is calculated in step e)
   by counting the number of immediate neighbors of the query word, whereby an
   immediate neighbor is a word that is connected through one link to the query word.

3. The method of claim 1, wherein the database (10) stores a plurality of texts (17).

4. The method of claim 1, comprising performing a search to find texts (11, 12, 13) in
   the database (10) that contain the query word.

5. The method of claim 4, wherein the steps b) through e) are repeated for each text (11,
   12, 13) that contains the query word.

6. The method of claim 5, comprising displaying a list (82) showing the characterizing
   strength (C) of each text (11, 12, 13) that contains the word.

7. The method according to any one of the preceding claims, wherein a parser is
   employed, to create the graph in step b).

8. The method of any one of claims 1 to 6, wherein a semantic network generator is employed to create the graph (30) in step b).

9. The method of any one of claims 1 to 3, wherein one graph is generated for each sentence in the text and wherein the characterizing strength (C) is calculated for each sentence by performing the steps b) through e).

10. The method of claim 9, wherein the characterizing strength (C) of the text is calculated in dependence on the characterizing strengths (C) of all sentences of the respective text.

11. The method of any one of claims 1 to 3 , wherein the graph is evolved in step c) by removing all words from the text that are not nouns and/or verbs.

12. The method of any one of claims 1 to 3, wherein the graph is evolved in step c) by replacing auxiliary verbs with main verbs.

13. The method of any one of claims 1 to 3, wherein the graph is evolved in step c) by leaving out verbs.

14. The method of any one of claims 1 to 3, wherein the subject of the sentence is identified and placed centrally in the graph to produce a tree-like graph structure in which the subject is at the root, prior to carrying out step d).

15. The method of claim 2, comprising the step of determining the number of second neighbors of the query word, whereby a second neighbor is a word that is connected through two links to the query word.

16. The method of claim 2 or 15, wherein the characterizing strength (C) of the text is an average calculated by
   • adding the characterizing strengths (C) of all sentences of the respective text, and
   • then dividing the result of the previous step by the number of sentences.

17. A system for automatically determining a characterizing strength (C) which indicates how well a text (17) in a database (10) describes a search query (15), the system comprising:

- o a database (10) storing a plurality of *m* texts (17);
- o a search engine (16) for processing a search query (15) in order to identify thoses *k* texts (11, 12, 13) from the plurality of *m* texts (17) that match the search query (15); and,
- o a calculation engine (18) for calcuating the characterizing strengths (C) of each of the *k* texts (11, 12, 13) that match the search query (15), by performing the following steps for each such text:
  - creating a graph with nodes and links, whereby words of the text are represented by the nodes and the relationship between words is represented by the links,
  - evolving the graph according to a pre-defined set of rules,
  - determining the neighborhood of the word, whereby the neighborhood comprises those nodes that are connected through one or more links to the word, and
  - calculating the characterizing strength (C) based on the topological structure of the neighborhood.

18. The system of claim 17, wherein the database (11) is stored in a server (90) connected via a network (94) to a client system (91, 92, 93).

19. The system of claim 17 comprising a parser for creating the graph.

20. The system of claim 17 comprising a semantic network generator for creating the graph.

21. The system of claim 17, wherein the calculation engine calculates the characterizing strength (C) by counting the number of immediate neighbors of the word, whereby an immediate neighbor is a word that is connected through one link to the word.

22. An information retrieval system comprising a system as claimed in any of claims 17 to 21.

23. A server computer system comprising a system as claimed in any of claims 17 to 21.

24. A client computer system comprising a system as claimed in any of claims 17 to 21.

25. Software module for automatically determining a characterizing strength (C) which indicates how well a text in a database describes a query, whereby said software module, when executed by a programmable data processing system, performs the steps:

a) enabling a user to define a query (15) comprising a word,

b) creating a graph (71) with nodes and links, whereby words of the text (17) are represented by nodes and the relationship between words is represented by means of the links,

c) evolving the graph (72) according to a pre-defined set of rules,

d) determining the neighborhood of the word, whereby the neighborhood comprises those nodes that are connected through one or a few links to the word, and

e) calculating the characterizing strength (C) based on the topological structure of the neighborhood;

f) displaying the characterizing strenth (C).

26. The software module of claim 30 comprising a search engine (16) for identifying those texts (11, 12, 13) in a plurality of texts (17) that match the query.